

Title: Assessment of the limits of neural phase-locking using mass potentials

Abbreviated Title: Phase-locking limits of the round window neurophonic

Authors: Eric Verschooten¹, Luis Robles^{1,2}, Philip X. Joris¹

¹ Laboratory of Auditory Neurophysiology, KU Leuven, Leuven, Belgium, Campus
Gasthuisberg O&N 2, Herestraat 49 bus 1021, B-3000 Leuven, Belgium

² Program of Physiology and Biophysics, Universidad de Chile, Santiago, Chile,
Independencia 1027, Santiago

Corresponding author: Philip X. Joris

Philip.Joris@med.kuleuven.be, Laboratory of Auditory Neurophysiology, Campus
Gasthuisberg O&N 2, Herestraat 49 bus 1021, B-3000 Leuven, Belgium, Tel : +32
(0)16 33 07 44

Email addresses remaining authors:

Eric Verschooten : eric.verschooten@med.kuleuven.be

Luis Robles : lrobles@med.uchile.cl

Number of pages : 30

Number of figures : 14

Number of words for Abstract : 149/250

Introduction : 507 /500

Discussion : 1447/1500

Conflict of Interest: The authors declare no competing financial interests.

Acknowledgements: Supported by BOF (OT/09/50) (Flanders, Belgium) and BIL11/12. We thank Dr. D. Johnson (Rice University, TX) for supplying nerve single fiber synchronization data (Fig. 14A).

Abstract

In the diverse mechanosensory systems that animals evolved, the waveform of stimuli can be encoded by phase-locking in spike trains of primary afferents. Coding of the fine-structure of sounds via phase-locking is thought to be critical for hearing. The upper frequency limit of phase-locking varies across species and is unknown in humans. We applied a method developed earlier, based on neural adaptation evoked by forward masking, to analyse mass potentials recorded on the cochlea and auditory nerve in the cat. The method allows us to separate neural phase-locking from receptor potentials. We find that the frequency limit of neural phase-locking obtained from mass potentials was very similar to that reported for individual auditory nerve fibers. The results suggest that this is a promising approach to examine neural phase-locking in humans with normal or impaired hearing or in other species for which direct recordings from primary afferents are not feasible.

Introduction

Neural phase-locking is a fundamental property of many mechanosensory systems. In auditory systems, it refers to the ability of neurons to synchronize their spikes to temporal features of the sound waveform. This is typically assessed with pure tones, to which neurons can discharge maximally at a preferred phase angle (Galambos and Davis, 1943; Tasaki, 1954; Kiang et al., 1965; Rose et al., 1967). Neural phase-locking has been proposed to be critical for many aspects of hearing and its disorders (reviewed in Lorenzi et al., 2006; Moore, 2008). Phase-locking in single fibers of the auditory nerve (AN) declines above a certain cut-off frequency and becomes undetectable above an upper frequency limit (Rose et al., 1967; Rose et al., 1968; Johnson, 1980; Palmer and Russell, 1986; Joris et al., 1994a), which differs between species (Weiss and Rose, 1988b). It has been most extensively studied in the cat (Bourk, 1976; Johnson, 1980; Rhode and Smith, 1986; Blackburn and Sachs, 1989; Joris et al., 1994b), using the vector strength (VS) metric (Goldberg and Brown, 1969), which decreases at a cut-off frequency of about 1 kHz and becomes insignificant above ~5 kHz.

The cut-off frequency and upper frequency limit of phase-locking in the AN of humans are unknown, because its traditional assessment requires intracranial access and penetrating microelectrodes. Psychophysical studies in human show a hard limit of <1.5 kHz up to which the binaural system can make use of stimulus fine-structure (Klumpp and Eady, 1956; Zwislocki and Feldman, 1956; Schiano et al., 1986; Brughera et al., 2013). This suggests that the limit of phase-locking to fine structure in human may be lower than in laboratory animals like the cat (Joris and Verschooten, 2013). However, other studies have argued that monaural hearing has access to fine-structure at much higher frequencies, perhaps up to 10 kHz (Heinz et al., 2001; Moore, 2008).

One specific form of mass potentials recorded on the AN, called the auditory nerve neurophonic (Snyder and Schreiner, 1984), provides a possible means to estimate and compare phase-locking across species, but this still requires intracranial access. Several studies reported a phase-locked neural component at the round window (RW) (Henry, 1995; He et al., 2012; Lichtenhan et al., 2013; Forgues et al., 2014; Verschooten and Joris, 2014). This component might be recordable in

humans using an approach through the middle ear (Eggermont, 1976), but it is heavily entangled with the cochlear microphonic (CM) from the transduction currents of hair cells (Dallos, 1973).

A previous paper from our laboratory documented a method to disambiguate and quantify the neural and hair cell components, based on forward masking and stimulus polarity reversals and focusing on a single frequency (Verschooten and Joris, 2014). Here, we apply the method over a broad frequency range in cat. The results show excellent agreement with single AN-fiber data and suggest that recording of mass potentials is a promising avenue to study peripheral phase-locking. It may enable an objective assessment of the proposed deficits in fine-structure coding in patients with impaired hearing (Starr et al., 1996; Lorenzi et al., 2006; Moore, 2008).

Materials and Methods

A detailed description of the methods can be found in a previous study (Verschooten and Joris, 2014). In what follows, we give a brief summary and additional details where needed.

Surgical Procedure

The experiments were conducted in 4 adult cats of either sex. All procedures were approved by the KU Leuven Ethics Committee for Animal Experiments and were in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals.

Animals were briefly examined for clear external and middle ears. Cats were anesthetized with a mixture of ketamine (20 mg/kg) and acepromazine (0.2 mg/kg) administered intramuscularly. A venous cannula was then placed to infuse lactated Ringer's solution and sodium pentobarbital to maintain a deep state of anesthesia. A cannula was inserted into the trachea. During the measurements the animals were kept warm with a homeothermic blanket (Harvard). The experiments were conducted in a double-walled soundproofed and faradized room (Industrial Acoustics Company).

One pinna was surgically removed and the auditory bulla exposed and opened. Through the opening in the bulla a silver, Teflon-insulated wire with ball electrode was inserted and placed at the round window. The electrode lead was glued to the bone and the bulla opening was closed again with ear impression compound (Microsonic). Two silver wire electrodes were threaded through the skin: a reference electrode at the nape of the neck and a ground electrode near the contralateral bulla. The AN was exposed *via* a posterior fossa approach, involving the removal of a small area of cerebellum. The AN measurements were performed with insulated Platinum/Iridium (90/10) ball electrodes mounted on a custom designed electrode holder for differential nerve recordings.

In two experiments we applied 4 μ L (10 mM) tetrodotoxin (TTX; Gentauro; powder dissolved in artificial CSF with citrate buffer, pH 4.8) in the niche of the RW to block neural spiking in the cochlea. Further details are given in Results.

Stimulus generation

Stimuli were generated with custom software and a digital sound system (Tucker-Davis Technologies, system 2, sample rate: 125 kHz/channel) consisting of a digital-to-analog converter (PD1), a digitally controlled attenuator (PA5), a headphone driver (HB7) and an electromagnetically shielded acoustic transducer (dynamic electro-acoustic transducer, Radio Shack, 20 Hz - 50 kHz). The transducer was connected with plastic tubing to a custom earpiece coupler. The coupler was fit in the transversely cut ear canal. The acoustic system was calibrated *in situ*, through the custom coupler and within a few mm from the eardrum, with a calibrated probe-microphone (Brüel & Kjær, type 4192, ½-inch condenser microphone and conditioning amplifier Nexus 2690).

Signal recording

We use the term “mass potentials” to refer to any of the potentials generated by a population of cells and recorded at some distance from these generators. Mass potentials are traditionally labeled with acronyms or names indicating specific stimuli and/or recording locations. We use the traditional names to the extent possible, and provide new names when needed for clarity. The potentials were measured with battery operated low-noise differential preamplifiers (Signal Recovery, Model 5115 and/or Stanford Research Systems, SR560) located in the shielded room. For the AN, different electrode configurations and positions were evaluated in a preliminary study. For the data reported here, the first two experiments were conducted with differential ball electrodes placed longitudinally on the nerve, with the active electrode closest to the internal auditory meatus. Two later experiments were conducted with a ‘monopolar’ configuration, with the active electrode placed on the auditory nerve and the reference at the nape of the neck. The RW measurements were performed between the active electrode at the round window and the reference electrode in the nape of the neck. The grounds of the preamplifiers were connected to the electrode at the contralateral mastoid. The signals were filtered (30 Hz - 30 kHz, cut-off slopes 12 dB/oct.) and further amplified with an external amplifier (Dagan, EX4-400) to a total gain of x10 000. All of the relevant signals including the electrical responses, the stimulus waveform, and synchronization pulses at the start of every probe were visualized on an oscilloscope (LeCroy, WaveSurfer 24Xs); moreover, the responses were sampled with an ADC (TDT, RX8, ~100 kHz/channel, max. SNR 96 dB) and stored on disk for further signal

processing. To increase the signal-to-noise ratio (SNR) of the response, the stimuli (positive polarity & negative polarity) were repeated at least 100 times and thereafter averaged. All responses shown in this study are averages.

Stimulus Paradigm

We used the paradigm described in Verschooten and Joris (2014) and illustrated in Figure 1. It was designed with an emphasize on the quantitative extraction of the entangled neurophonic and the removal of artifacts. The complete sequence consisted of 3 segments given at one stimulus polarity, alternated with the identical stimuli given at the opposite polarity. The first segment (a) contains only a pure tone probe stimulus; the second segment (b) the same probe but now preceded by a masker with a separation of 1 ms; and the last segment (c) only the masker used in (b). The probes had durations of 50, 100 or 150 ms, depending on the experiment. The maskers were either pure tones at the probe frequency (f_p), or a fixed broadband Gaussian noise (typically 50-8000 Hz) with a fixed duration of 100 ms (illustrated in Fig. 1). For the purposes of the present study, the nature of the masker (pure tone or broadband noise) is irrelevant. The interval between different segments was at least 50 ms. Segment c was designed to have a total length equal to or greater than segment b, and provides a large stimulus-free period. To reduce spectral splatter at the stimulus transients, the probe and masker were gated with a 1 ms raised-cosine. The probe level (50 or 55 dB SPL) was chosen such that the SNR of the probe response and the number of presentations required were adequate and realistic; and the masker level (65 or 70 dB SPL) was chosen such that the probe response was properly masked. For broadband noise maskers, the noise masker level was calculated over the total noise bandwidth.

Analysis

The responses of the segments in Figure 1 were combined and processed to obtain the desired signals. One full stimulus sequence resulted in 2 *pairs* of responses. Figure 2 is an example of these recorded responses and their derived signals at the AN and RW for a probe frequency of 1 kHz. The two top pairs of responses (Fig. 2a,b) are the probe and masked probe responses, for stimuli of opposite polarity (cf. Fig. 1a,b). From these pairs of responses, a purely neural signal is obtained by

subtracting the masked response (Fig. 2b) from the probe response (Fig. 2a): we refer to this difference as the adapted component (Fig. 2c). This difference is purely neural because adaptation is present in AN-fibers and not in receptor potentials (Russell and Sellick, 1983; Palmer and Russell, 1986; Eatock, 2000); but it contains both phase-locked and non-phase-locked neural components. For the purposes of the present study, we are particularly interested in the phase-locked component. The non-phase-locked component is the so-called compound action potential (CAP) (Goldstein and Kiang, 1958; Kiang et al., 1976; Antoli-Candela and Kiang, 1978), which reflects the synchronous firing of AN-fibers to stimulus onset. It was removed from the adapted component by filtering (non-causal phase preserving high-pass Chebyshev FIR filter at $\frac{1}{2}f_p$), which results in the signal of interest which we term the decaying neurophonic (dNP; AN-dNP for the AN and RW-dNP for the RW). This signal is shown in Figure 2d and has been argued to contain only phase-locked neural contributions (Verschooten and Joris, 2014). Note that the “neural purity” of this signal does not depend on complete masking: if some neural signal is left in the masked responses, the adapted components will be smaller and in the worst case it will be non-existing, but this will not introduce CM contamination into the adapted components and other signals derived from them.

The availability of responses to stimuli with opposite polarity allows us to further dissect the adapted components. From a pair of decaying neurophonic traces (Fig. 2d) we derived two other signals: the stimulus polarity dependent response (Fig. 2e) and the stimulus polarity independent response (Fig. 2f), by calculating their halved difference (obtained from (magenta trace - cyan trace)/2 in Fig. 2d) and their halved sum (Fig. 2f). To the latter sum we applied the same filter as for trace 2d, but with a high pass cut-off at f_p rather than at $\frac{1}{2}f_p$. Due to the coherent phase relations within the pairs of the adapted components, the phase-locked odd harmonics (typically dominated by the 1st harmonic) and even harmonics (typically dominated by the 2nd harmonic) are separated into the stimulus polarity dependent response and stimulus polarity independent response, respectively. The time course and the maximum magnitudes of the various dNP signals (cf. Fig. 2d-f) and their harmonics (1st, 2nd) or combined harmonics (1st-4th) are derived with an Hilbert transformation (Fig. 3-6) and a short-time Fourier transform (STFT; here implemented as a GABOR transform with a time window of 6 cycles of f_p) (Fig. 6-14), respectively. For some signals, we also measured *steady-state amplitudes* (Fig. 6 and

10), which were measured over the stable and transient-free period of the response using a Fast Fourier Transformation.

Noise floor. For the assessment of the upper frequency limit of phase-locking, the noise floor of the derived signal was quantified. It was calculated from the background noise in the stimulus and artifact free part (Fig. 1, segment c, 540 - 706 ms) taking into account the different response combinations and operations (e.g. subtraction, STFT, etc.) that were used to quantify the corresponding responses. Segment c served also another function: in the response of segment b there is an offset response to the masker, which interferes with the probe response. To remove this off-set response we always removed this response in segment b by subtracting the response after the masker in segment c from that in segment 2. More details about the different extraction methods can be found in (Verschooten and Joris, 2014). Except when mentioned otherwise, the amplitudes are presented as a peak voltage.

Results

We investigated the frequency dependence of auditory neural phase-locking in cat using a method (Verschooten and Joris, 2014) that is able to isolate and extract a measure of the neurophonic at the RW. We verified the method over a wide frequency range (300 Hz - 8 kHz) with simultaneous recordings at the AN and by applying a neural blocker (TTX). We illustrate our findings with examples from individual animals, and give a population analysis in the final section. We derive the upper frequency limit of neural phase-locking from these mass potentials and discuss how the results relate to the phase-locking limit of single AN-fibers.

Frequency dependence.

Figure 3 presents the time course of the decaying auditory neurophonic for a wide range of probe frequencies at the AN (panel A) and RW (panel B). These curves correspond to the envelopes illustrated with thick lines in Figure 2d. A first feature of note is that the amplitudes are higher when recorded at the RW than on the AN (Verschooten and Joris, 2014). Second, we observe that, regardless of frequency and measurement location, the time course of the envelopes is always similar,

showing an early peak followed by an exponential decay. For frequencies with low SNR, the tail of the decay was not distinguishable from the background noise; in some cases only the region around the peak remained. Third, two distinct features were frequency dependent: the (peak-) magnitude and the delay of the rising slope. Peak-magnitude increases with increasing frequency but decreases again above 1 to 1.2 kHz. The delay of the rising slope decreased with increasing probe frequency, at least up to a few kHz, after which the magnitude becomes too small to make meaningful statements regarding delay. The peak-magnitudes of Figure 3 are graphed in Figure 4A,B as a function of probe frequency. The lower panels (Fig. 4C,D) show similar measurements, from the same animal, using tonal maskers. All curves show the same trend: magnitude increases with frequency, reaches a maximum, then decreases sharply until it is bounded by the noise floor. The noise floor is indicated by the dashed line and is the baseline of the neurophonic (see Methods). The detectable upper frequency limit is defined here simply as the frequency where the curve reaches the noise floor (we return to this definition in Fig. 12). With this definition, and using a broadband masker, the limit is 5.0 kHz at the AN but was not reached at the RW. Using a tonal masker, the limits were 5.4 kHz and 5 kHz at the AN and RW, respectively. We visually mark an approximate center frequency for the broad peak of the general band-pass characteristic in these figures. For the two measurement sessions of Figure 4, these center frequencies were quite similar for the AN and RW and were between 1.1 and 1.5 kHz.

In broad terms, the characteristics of Figure 4 show a similarity to limits of phase-locking in auditory nerve fibers (Rose et al., 1967; Johnson, 1980). This observation, combined with the fact that the general shape of these characteristics is very similar for the two recording locations (AN vs. RW), holds promise to estimate neural phase-locking from mass potential recordings at the middle ear. In the following sections (Figs. 6, 9 and 10), we perform a finer analysis to critically examine whether the dNP indeed reflects a purely neural signal.

Harmonic separation. We first extracted the peak-magnitudes of the maskable odd harmonics (dominated by the 1st) and even harmonics (dominated by 2nd) using the decomposition of the subtracted, i.e. stimulus polarity-dependent, responses (e.g. Fig. 2Ae,Be) and of the summed, i.e. stimulus polarity-independent, responses (e.g. Fig. 2Af,Bf). Figure 5 shows the odd (red triangles) and even (blue squares) harmonic contributions to the AN-dNP and RW-dNP peak-magnitudes. Below

600 Hz, the amplitudes of the odd and even harmonics are comparable. Above 600 Hz, the total neurophonic response is clearly dominated by the odd harmonics; the trace of the even harmonics has a similar course but is up to 20 dB (i.e. a factor of 10 of the voltage scale) lower. As a consequence, the maximum detectable frequency limit of the even harmonic response is well below that of the odd harmonic response, which in its turn has a slightly lower detectable limit than the total response (black, same data as in Fig. 4). Similar trends were found in other experiments.

For two reasons, we also wished to go beyond these decomposed signals and to more purely quantify the 1st and 2nd (rather than the entire uneven and even) harmonic content. Typical measures for phase-locking in single AN-fibers are based on the 1st harmonic of the neural response and therefore not on odd, even or all harmonics such as the peak-magnitude. Also, in studies using mass potentials, the 2nd harmonic is commonly used to differentiate neural from hair cell contributions (Snyder and Schreiner, 1984; Lichtenhan et al., 2013). In a procedure that we refer to as the STFT method, we used a Gabor transform to derive the maximum amplitudes of the 1st and 2nd harmonics from the decomposed signals obtained in the first step. Note that these harmonics can also be extracted directly from the original, non-decomposed paired dNP signal (e.g. Fig. 2Ad), but the decomposition has the advantage to easily remove the CAP in the stimulus polarity dependent response, and to favor the phase-coherent parts in the harmonics.

The results of the STFT extraction for the tonal maskers in Figure 5 are shown in Figure 6: the 1st harmonic (solid red, Fig. 6A, B) and 2nd harmonic (solid blue, Fig. 6C, D) are shown together with the polarity-dependent and -independent peak-amplitudes, replotted from Figure 5 as the dashed traces. Above 600 Hz, the traces of the different harmonics have a similar frequency dependent course as their related basic signals (dashed lines) but ~6 dB lower in amplitude. The upper frequency limit with respect to the noise floor is comparable to those of the peak-amplitudes (Fig. 5). Below 600 Hz, the course of the 1st and 2nd harmonic deviate significantly from the peak-amplitude of their basic signals: the magnitudes are much lower with a more pronounced high-pass characteristic. To understand this difference, and to address the validity of our method across frequency, we compared the STFT 2nd harmonic (Fig. 6C,D, solid blue line) and peak-amplitude (Fig. 6C,D, dashed blue line) of the stimulus polarity independent responses with the 2nd harmonic steady-state amplitude in the

stimulus polarity independent probe response (Fig. 6C,D, dot dashed blue line; cf. the stable amplitude of the sustained part of the halved summed responses in Figure 2Aa, Ba, sometimes called the ANOW (Lichtenhan et al., 2013; Forgues et al., 2014)). The reasoning is that if this 2nd harmonic steady-state component is purely neural, as has been argued in other studies (Snyder and Schreiner, 1984; Lichtenhan et al., 2013; Forgues et al., 2014), then it provides a means to verify the frequency dependent course of the 2nd harmonic of the STFT. Notwithstanding a different magnitude and noise floor (not shown for the peak-magnitude), the STFT (solid blue) and the peak-magnitude (dashed blue) traces show similarities to the steady-state response (dashed dotted, cyan): a broad best frequency region and a high-frequency fall-off, with frequency limits (vertical dashed lines) that are in reasonable agreement. The main difference is below 600 Hz: the 2nd harmonic steady-state and the STFT traces do not show the plateau seen in the peak-magnitudes, but both have a high pass characteristic which suggests a mutual relationship. As a result, below 600 Hz the peak-magnitude of the polarity-independent responses, and most likely also for the dependent responses, is not a reliable measure for the neurophonic. Further analysis (not shown), revealed that the deviation between the trace of the peak-magnitude and the other traces (STFT and steady-state) reflects leakage of the CAP: at low frequencies part of the power spectrum of the CAP appears in the calculated responses, despite the polarity reversal technique and the CAP-removal filter. Note that the higher magnitude for the STFT and peak-magnitude simply reflects the larger amplitude at neural onset, whereas the level difference in noise floor is due to the length of temporal integration, which is much longer for the steady-state measurement than for the STFT and peak-magnitude measurements.

Tonal versus noise maskers. As mentioned earlier, we used two types of maskers: a fixed broadband noise masker, to mask neural activity over a broad frequency range, and tonal maskers with the same frequency as the probe. We compared the results for the two types of maskers and did not find any substantial differences at the frequencies of interest, only some small deviations at the lowest probe frequencies. This is illustrated in Figure 5 for the peak-magnitudes, and illustrated in Figure 7 (solid versus dashed traces) for the combined first four harmonic contributions and the 1st and 2nd harmonics obtained with the STFT. Importantly, the upper frequency limit is basically independent of masker type - as expected.

TTX

As a final test of the neural basis of the components that we designate as such, we compared results obtained before and after the application of Tetrodotoxin (TTX) at the RW in two animals. The results were similar in the two animals and are illustrated for one animal. TTX blocks the firing of action potentials in neurons, but does not affect the ionic transduction currents through the hair cells, the synaptic currents in the dendrites of the auditory nerve fibers and the summing potential (extracellular DC potential of the hair cells), which is filtered by the presence of the high-pass filter (Materials and Methods). We used the neural blocker to measure the RW-CM in the absence of neural spiking, and to test the effectiveness of our method to cancel the RW-CM.

Effect on CAP. After an initial “normal” measurement session to obtain baseline estimates of the AN-dNP and RW-dNP, we applied 4 μ L (10 mM) TTX at the niche of the RW and monitored the averaged ($n = 100$) CAPs (maximum peak value minus peak N1) as the TTX took effect. The stimuli used were ON/OFF-gated non steady-state tone pips with a duration of 10ms and a sound level of 60 dB SPL. The evolution of the magnitude of the CAPs is shown in Figure 8A for 8 probe frequencies ranging from 1 to 10 kHz. The post-TTX time at which CAP reduction starts, differs at different frequencies, and these times are consistent with progressing diffusion of TTX through the cochlea. The CAP for high frequencies, which excite the most basal receptors, was affected first, followed by the CAP in response to lower frequencies. Remarkably, for frequencies at 3 kHz and below, there was a period of enhancement where the CAPs grew above its initial magnitude. The CAP then declined simultaneously at these frequencies.

The criterion to start acquisition of a new set of neurophonic measurements was a stable CAP reduction for all measured frequencies. For the animal illustrated in Figure 8, this was 4 hours after the administration of the TTX, as indicated by the dashed line. Note that at the time of neurophonic measurement (gray zone in Fig. 8A), there is a remaining CAP residue which is highest for mid-frequencies. This residue is thought to arise from the excitatory post-synaptic potentials of the afferent dendrites to the inner hair cells (Dolan et al., 1989). After an additional time period of 2 hours over which the neurophonic measurements were completed, the CAP was measured again: some small

additional magnitude decline was measured at mid-frequencies. A further decrease was noted at a final repeated measurement 18 hours after TTX application (Fig. 8, 1080 minutes).

Effect on CM. The decrease in CAPs does not simply reflect a loss of cochlear sensitivity. Figure 8B illustrates the effect of TTX on the 1st harmonic of the RW-CM for a wide range of frequencies. The magnitude of the pre-TTX RW-CM (blue trace) was measured on a short initial segment of the masked stimulus polarity dependent response just before neural onset, over which only CM response is present (cf. Fig. 2, the duration between onsets of masked and adapted components) (Verschooten and Joris, 2014). The magnitude of the post-TTX RW-CM (red trace) was obtained at steady-state using an FFT; its noise-floor, calculated from a stimulus-free part of the response, is shown as the dotted line with shading. The magnitude of this post-TTX steady-state signal was virtually identical to that measured at the response onset (not shown), which testifies to the validity of the onset-CM values as a reliable measure of the hair-cell generator. The pre- and post-TTX RW-CM magnitudes are remarkably similar (Fig. 8B): the only noticeable difference is a small decrease in magnitude of only a few dB, and this over the whole frequency range. At low frequencies, both RW-CM traces show a high-pass behavior with a slope of 40 dB/decade. Above 1.2 kHz the spectrum has little overall slope but shows a spectral fine structure with distinct valleys and peaks. The similarity between the two curves is consistent with a purely neural effect of TTX. The small amplitude decrease and the presence of the peaks and valleys in the fine structure are consistent with previous studies (Henry, 1995; He et al., 2012) if it is taken into account that these studies did not attempt to differentiate the neural and non-neural contributions measured at the RW. Taken together, the results indicate that the RW-CM is of the same origin and same magnitude before and after TTX application, and that it contains no significant neural activity.

Effect on neurophonic. Figure 9 shows the magnitudes of the AN-dNP (left column) and RW-dNP (right column) before (upper row) and after (lower row) TTX administration. Comparisons between the upper and lower rows show that TTX has a large suppressive effect on the neurophonic at both locations. The total neurophonic (black trace, combination of the first four harmonics) at the AN was reduced by more than 20 dB and became indistinguishable from the background noise. At the RW, the reduction was 26 dB, but some components in the total neurophonic, more specifically the 1st

harmonic (red trace) remained present between 0.4 and 2 kHz. Since this frequency range corresponds to the broad peak of the neurophonic (panel B), the residue is almost certainly of neural origin and indicates that the neural block was incomplete, as supported by the remaining post-TTX CAP in Figure 8A. This could not be verified with the results at the AN because of the lower SNR at that location. Remarkably, the noise floor at the RW decreased much more than at the AN. This reduction is indicated in Figure 9D by the gray arrows between the noise floor before (gray dotted curve) and after (black dotted curve) TTX. The decrease in noise floor for the RW was ~ 11 dB, while it was only ~ 1.3 dB for the AN (Fig. 9C). In contrast to the AN, the noise floor at the RW was reshaped from a band-pass shape to a straight line. The slope of this straight line is $\sim +7$ dB/decade, which is close to the slope ($+10$ dB/decade) imposed by the frequency proportional filter in our method. The change in spectrum of the RW noise floor to a flat slope close to 10 dB/decade suggests that the background noise of the cochlea became whiter, which indicates that the (spontaneous) neural background spiking activity is quenched (Dolan et al., 1990; McMahon and Patuzzi, 2002; Patuzzi et al., 2004; Searchfield et al., 2004). It is unclear why the noise floor at the AN did not reduce with the same amount as at the RW. One possibility is that the background noise at the AN only partially reflects spontaneous neural AN activity. The measurements at the AN in this experiment were recorded with respect to the nape of the neck and not longitudinal over the nerve as in some of the other experiments. Since the AN is not electrically isolated from neighboring structures such as the vestibular nerve, the noise floor is perhaps dominated by generators other than AN-fibers that are not affected by the TTX.

Steady state and 2nd harmonic. The steady-state 2nd harmonic measured in the portion of stimulus polarity independent response (cf. Fig. 6C, D) has been used as a measure of neural phase-locking (Lichtenhan et al., 2013; Forgues et al., 2014). Having shown that the RW-CM after TTX contains little neural contribution (Fig. 9), we can examine whether this steady-state 2nd harmonic effectively measures neural phase-locking, and whether it can be used to find its upper frequency limit.

Figure 10A shows the steady-state 2nd harmonic amplitude at the RW pre-TTX (blue) and post-TTX (red). Note that the steady-state pre-TTX trace is reasonably similar to that of the 2nd harmonic RW-dNP illustrated in Figure 9B (blue line), but that it shows several additional peaks

above 2 kHz. Similar peaks are present in the 1st harmonic of the pre- and post-TTX RW-CM (Fig. 8B), suggesting that these peaks are of receptor rather than of neural origin. The post-TTX response and noise floor (Fig. 10A, red) are clearly reduced for frequencies <2 kHz, up to 30 dB. This indicates that there is a strong contribution of neural activity to the steady-state 2nd harmonic at these frequencies, but on the other hand a large signal remains even after TTX, unlike the 2nd harmonic component of the RW-dNP (Fig. 9, blue) that disappears into the background noise. Also, above 2 kHz the post-TTX response is only reduced by a few dB, if not bounded by the noise floor (red dotted line). This limited reduction, combined with the observations that the peaks above 2 kHz are similar to those in the 1st harmonic RW-CM (red and blue traces, Fig. 8B) and are not present in the 2nd harmonic of the RW-dNP (blue trace, Fig. 9B), strongly suggests that these remaining contributions are dominated by CM. According to this interpretation, the 2nd harmonic steady-state response, (ANOW, Lichtenhan et al., 2013), contains a hair cell contribution which is relative large at high frequencies, which makes it unsuitable to estimate neural phase-locking at high frequencies.

A caveat is that TTX did not completely block all neural responses (see RESULTS, Fig. 8A and Fig. 9D), so there is still a possibility that the remaining steady-state response (Fig. 10A, red line) below 2 kHz is not a contamination by CM but a remaining non suppressed neural component. To investigate this, we overlay post-TTX 2nd harmonic amplitudes at steady state (red trace, same trace as in panel A) with those from the RW-dNP (green trace, same as blue trace in Fig. 9D). The reasoning is that the latter trace should be completely devoid of a neural contribution, because of both TTX and forward masking. Figure 10B shows that there is no trace of a response over the complete probe range in the post-TTX RW-dNP, not even in the region between 800 and 1300 Hz where the steady-state amplitude is larger. Note that the 2nd harmonic steady-state amplitudes are typically not larger than those of the RW-dNP (see Fig. 6C, D): thus if the remaining steady-state post-TTX response (Fig. 10A, red line) were neural in origin, it is expected to also be clearly present in the RW-dNP. The conclusion is that the steady-state 2nd harmonic is contaminated by CM, where this is clearly not the case for the RW-dNP. To summarize, the steady state 2nd harmonic signal at the RW is not suitable for determination of the upper frequency limit of phase-locking, because of the presence of a non-maskable component which probably arises from the CM.

Population analysis

In previous sections, we focused primarily on the method and its verification with some representative examples. In this section we combine the entire data set that is obtained with the STFT-method and determine an overall upper frequency limit of phase-locking. Figure 11 shows the single data points of 6 sessions for the various harmonics of the AN (panel A) and RW (panel B). Shown are the magnitude of the dNP (black) and its 1st (red) and 2nd (blue) harmonic. Trend lines (MATLAB, *RLOESS*, span: 0.35) use the same color code. For both locations, the different trend lines show a band-pass shape with a center frequency around 1 kHz. At this center frequency and above it, the 1st harmonic is the dominant component; below 1 kHz, the 2nd harmonic has a significant contribution which equals that of the 1st harmonic at frequencies < 500 Hz; and at the lowest frequencies (200 and 300Hz) even more higher harmonics contribute. We remark that responses to probe frequencies above 7 kHz were measured in only one experiment, which biased the trend line toward the values of that experiment. Therefore, only the relevant part up to 6 kHz is shown in Figure 11.

Determination of an upper frequency limit requires a baseline or noise floor. The trends of the noise floors were calculated in the same way as for the neurophonic and are included in Figure 11 (dashed lines; single data points not shown). The trend lines for the 2nd harmonics touch the accompanying noise floor. This is not the case for the other trend lines, where at about 5 dB above the noise floor the curves start to bend towards the noise floor but remain ~1 dB above these. This phenomenon could be due to a systematic bias in the calculation of the noise floor consequent to the 99.75% percentile noise amplitude criterion used (see METHODS). However in most of the above figures, an upper limit is reached above which there are multiple crossings between signal and noise floor. Another, more likely possibility is that this phenomenon results from a few curves that have not settled yet at the highest frequencies, e.g., Figure 5B. Whatever its cause, the fact that the signal remains at a constant level of ~1 dB above the noise floor indicates that at that point the noise floor has overtaken the response. To determine an upper cutoff frequency we therefore first simply raised the noise floors (by 0.2–1.5 dB) till they merged with the high-frequency part of the trend lines, where both curves run parallel. Next, we subtracted (power wise) these noise floors from the respective trend

lines, in the assumption that the noise and phase-locked signals are not correlated. The result of this subtraction is depicted in Figure 12 by the solid lines which have a decreasing slope extending up to $-\infty$ dB (= the imposed intersection between the noise floor and the original trend); the curves were truncated at -30 dB and are called the (noise) compensated trend lines. We then defined the upper frequency limit at 10 dB below the intersection of raised noise floor and compensated trend lines; this 10 dB point is indicated by the small horizontal lines. At the frequency at which the noise floor and compensated trend line intersect, illustrated for one case with a white dot (Fig. 12B), the SNR for the original trend line is theoretically 3 dB. The choice of the 10 dB corresponds to an SNR of ~ 1.7 dB. Using our new definition we obtained a mean detectable upper frequency limit of phase-locking for the RW at 5.4 kHz and 4.2 kHz for the AN. The limit for the 1st harmonic for the RW was at 4.7 kHz and at 4.2 kHz for the AN. The 2nd harmonic had in general a much lower limit, 2.2 kHz for the RW and 1.8 kHz for the AN.

AN-CM. As is clear from many figures reported in this study, the neurophonic is strongly attenuated at higher frequencies, consistent with the steep low-pass characteristic of single-fiber neural phase-locking (Weiss and Rose, 1988b, a; Kidd and Weiss, 1990). This slope is at least 40 dB/decade (see also later Fig. 14) and is much less steep for the CM (Henry, 1995; He et al., 2012). It follows that the AN-CM will inevitably dominate over the AN-dNP above a certain frequency. This is illustrated in Figure 13 where population trends of the AN-CM and AN-dNP are shown. The measurements are separated in two groups according to the recording method; recordings measured differentially along the nerve (solid lines) and those measured ‘monopolar’ with respect to the nape of the neck (dashed lines). The AN-CM trends (green curves) were similar for the two methods: they increase up to 0.8-1 kHz, thereafter they decline with a roll-off of -20 dB/decade and finally they intersect the neurophonic at a frequency of 3 kHz for the ‘monopolar’ recordings and at 4 kHz for the differential recordings. Beyond these frequencies, the AN-CM is larger than the AN-dNP. The AN-dNP trends (red lines) also show a band-pass trend centered near 1 kHz, but note that their low-pass slope has a distinctly steeper roll-off; it is roughly -50 dB/decade.

Discussion

Human behavioral results have been interpreted as indicating a crucial role for neural phase-locking to the fine-structure of sound waveforms, even at frequencies above the limit of neural phase-locking of commonly studied mammals (Heinz et al., 2001; Moore, 2008). However, the extent of neural phase-locking in human is debated (Joris and Verschooten, 2013) and calls for physiological measurements. Here we assess neural phase-locking with mass potentials recorded at the cochlear RW, in a species for which extensive single AN-fiber data are available. Previous authors have examined whether limits of neural phase-locking may be gleaned from mass-potential recordings. (Snyder and Schreiner, 1984) reported neurophonic recordings in cat that suggested neural phase-locking exceeding the generally quoted limit of ~5 kHz; on the other hand (Lichtenhan et al., 2013) expressed skepticism in the ability of such recordings to be a valuable tool at high frequencies. We used a method developed earlier (Verschooten and Joris, 2014) based on forward masking, polarity reversal and filtering, to separate the neural phase-locked contribution in the mass potential at the RW from the phase-locked CM. We examined neural phase-locking over a wide frequency range and obtained good agreement with the upper limit of neural phase-locking of single AN-fibers. The method can be adapted for clinical use, e.g., to detect the abnormal neural phase-locking hypothesized in auditory neuropathy (Starr et al., 1996).

Frequency limit at the round window

We show that RW mass-potentials have a considerable neural phase-locked contribution that can be isolated and quantified over a wide frequency range. This component, the RW-neurophonic, has a spectral band-pass shape with a center frequency around ~1 kHz. It is strikingly similar to that recorded at the AN, but with a typically larger magnitude (e.g., Figs. 3-7,9,11,12). The upper frequency limit of neural phase-locking of the 1st harmonic obtained at the RW was 4.7 kHz. This limit corresponds rather well with that reported for individual auditory nerve fibers in cat (~5 kHz, Johnson, 1980; Rhode and Smith, 1986; Javel and Mott, 1988; Joris et al., 1994a; van der Heijden and Joris, 2006).

Neural rectification introduces higher harmonics. The steady-state 2nd harmonic of the neurophonic is easy to extract and has been used as an indicator of neural phase-locking (Henry, 1995; Lichtenhan et al., 2013; Forgues et al., 2014). We examined this component and find that it is not suitable for the determination of the frequency limit of phase-locking. First, the limit is lower than for the 1st harmonic (e.g., Fig. 12). Second, CM contamination limits detection of small neural signals at high frequencies (Fig. 10), which is supported by a recent study in gerbil (Forgues et al., 2014).

Frequency limit at the auditory nerve

The mass potential recorded at the AN showed an upper frequency limit of phase-locking of 4.2 kHz. This is comparable to the limit (~4 kHz) found in an earlier study in cat (Snyder and Schreiner, 1984, their Figure 12) obtained from a transfer function at a level of 60 dB SPL. However, in that same study, transfer functions at higher levels (70, 80 dB SPL) indicated phase-locking extending to the highest frequencies tested (10 kHz), i.e., an octave above the limit in single AN-fibers (~5 kHz). In Figure 13 we illustrate that above a few kHz the stray-CM will inevitably dominate the neurophonic at the AN. This is to be expected from the steep low-pass characteristic (≥ 40 dB/decade) of single-fiber neural phase-locking (Weiss and Rose, 1988b, a; Kidd and Weiss, 1990). The transfer functions at higher probe levels in Snyder and Schreiner (1984) had a much less steep roll-off of -20 dB/decade compared to -40 dB/decade at 60 dB SPL and -50 dB/decade obtained here; but the -20 dB/decade roll-off is similar to that for the AN-CM in Figure 13. We thus surmise that the transfer functions in Snyder and Schreiner (1984) are contaminated by stray-CM, and that the degree of contamination depends on recording method, frequency and level.

The plateau of the peak-magnitude below 600 Hz

Figure 6 shows a clear amplitude plateau below 600 Hz for the peak-magnitudes (dashed), which was not present for the steady-state (blue dot-dashed) or for the STFT (solid). Peak-magnitudes contain higher harmonics besides the dominant harmonics (first or second). These higher harmonics are more present at low frequencies (e.g., Fig. 7B,C,E,F and Fig. 9A,B), but they cannot explain the amplitude plateau. Upon further investigation we found that the plateau reflects spectral leakage of the CAP. For

the polarity independent response (dashed curve, Fig. 6C,D) this was not unexpected because below 600 Hz the CAP's spectral energy is in the pass-band of the high-pass filter (cut-off frequency = f_p). For the polarity dependent response (dashed curve, Fig. 6A,B) the plateau is due to a slight timing difference between low frequency (<600 Hz) CAPs to opposite polarities; this is not further investigated here. This plateau is not present when using the STFT method (Fig. 6), due to its larger temporal integration and frequency selectivity. Above 600 Hz, the use of the STFT is not critical but desirable because it represents the 1st harmonic (cf. VS).

Relationship to phase-locking in single AN-fibers

An important question is how our results relate to the synchronization of single AN-fibers. A number of factors hamper direct comparison between maximal VS of single AN-fibers in cat (Johnson, 1980; Rhode and Smith, 1986; Javel and Mott, 1988; Joris et al., 1994b) and the neurophonic. First, the neurophonic is an ensemble response from many AN-fibers, with their own phase, delay, rate and VS depending on frequency, CF, and stimulus level (Rose et al., 1967; Anderson et al., 1971; Kim and Molnar, 1979; Palmer and Russell, 1986; van der Heijden and Joris, 2006; Palmer and Shackleton, 2009; Temchin and Ruggero, 2010). Spatiotemporal summation of these single AN-fiber responses is expected to affect the amplitude and phase of the recorded response in ways that are ill-understood. For example, the shapes of the transfer functions of the neurophonic in this and previous studies (Fig. 12 in Snyder and Schreiner (1984); Fig. 9 Henry (1995)) often show interference patterns, especially in the sharp declining slope, which are not present in population plots of maximal VS in the cat AN. These interference patterns are also visible in the microphonic (Fig. 13 and Fig. 14A). Second, whereas the VS of AN-fibers is based on histograms of the timing of action potentials, the neurophonic is an analog waveform composed of summed single unit responses (Kiang et al., 1976; Prijs, 1986; Versnel et al., 1992) whose waveform has a convolving low-pass filtering effect. Third, VS or synchronization index is a normalized metric which equals the Fourier magnitude at the frequency of interest divided by the overall mean spike rate of AN-fibers.

In an attempt to compare the neurophonic with single AN-fibers we can compute a synchronization index for the neurophonic (SIN) where the RW-CM is used for normalization. The

underlying assumption is that the RW-CM can be regarded as an input signal while the RW-dNP is the output signal, and furthermore that both signals undergo similar extracellular spatiotemporal filtering. The mean trend of the RW-CM is shown as the green curve in Figure 14A; the mean trend of the 1st harmonic component of the neurophonic is shown as the solid red curve. To compensate the neurophonic for the convolving effect of the single-unit response waveforms, we extracted an average normalized spectrum from four unit responses in cat (unit 9, 24, 45 and 43 of Figure 6 in (Kiang et al., 1976)). This spectrum is represented as the blue curve in Figure 14A and exhibits a shallow band-pass characteristic. The red dashed line is the 1st harmonic component of the neurophonic compensated for this spectrum. Finally, the brown line is the SIN for the neurophonic, calculated as the ratio between the compensated neurophonic (red dashed curve) and the RW-CM (green curve). Only the spectral shape of the resulting curve is of interest: its maximum is put at 0 dB. We compare it to the maximal VS in dB for single AN-fibers in cat (circles and black line; from Johnson (1980)). The SIN (brown) does not show the same smooth course as the VS: it fluctuates between 0 and -10 dB and declines at a higher frequency (3.3 kHz) than the VS (~2.5 kHz).

Our main interest is in the overall slopes of the trends. Figure 14B shows a speculative and idealized summary of the different curves shown in Figure 14A, with slope values in dB/decade. It shows a reasonably coherent picture of the mutual relationships between the different measurements. The neurophonic SIN shows a roll-off around 3 kHz with a slope of -60 dB/ decade, reasonably consistent with the trend line for VS in single fibers.

Reference

- Anderson DJ, Rose JE, Hind JE, Brugge JF (1971) Temporal Position of Discharges in Single Auditory Nerve Fibers within the Cycle of a Sine-Wave Stimulus: Frequency and Intensity Effects. *J Acoust Soc Am* 49:1131-1139.
- Antoli-Candela EJ, Kiang NYS (1978) Unit activity underlying the N1 potential. In: *Evoked Electrical Activity in the Auditory Nervous System* (Naunton R, Fernandez C, eds), pp 165-189. New York: Academic Press.
- Blackburn CC, Sachs MB (1989) Classification of Unit Types in the Anteroventral Cochlear Nucleus: PST Histograms and Regularity Analysis. *Journal of Neurophysiology* 62:1303-1329.
- Bourk TR (1976) In: *Electrical Responses of Neural Units in the Anteroventral Cochlear Nucleus of the Cat*, p 385. Ph.D. thesis: MIT.
- Brughera A, Dunai L, Hartmann WM (2013) Human interaural time difference thresholds for sine tones: the high-frequency limit. *The Journal of the Acoustical Society of America* 133:2839-2855.
- Dallos P (1973) *The auditory periphery*. New York: Academic Press.
- Dolan DF, Xi L, Nuttall AL (1989) Characterization of an EPSP-like potential recorded remotely from the round window. *J Acoust Soc Am* 86:2167-2171.
- Dolan DF, Nuttall AL, Avinash G (1990) Asynchronous neural activity recorded from the round window. *The Journal of the Acoustical Society of America* 87:2621-2627.
- Eatock RA (2000) Adaptation in hair cells. *Annu Rev Neurosci* 23:285-314.
- Eggermont JJ (1976) Electrophysiological study of the normal and pathological human cochlea. II. Neural responses. *Rev Laryngol Otol Rhinol (Bord)* 97 Suppl:497-506.
- Forgues M, Koehn HA, Dunnon AK, Pulver SH, Buchman CA, Adunka OF, Fitzpatrick DC (2014) Distinguishing hair cell from neural potentials recorded at the round window. *Journal of neurophysiology* 111:580-593.
- Galambos R, Davis H (1943) The response of single auditory nerve fibers to acoustic stimulation. *Journal of neurophysiology* 6:39-58.

- 565 Goldberg JM, Brown PB (1969) Response of Binaural Neurons of Dog Superior Olivary Complex to
 566 Dichotic Tonal Stimuli: Some Physiological Mechanisms of Sound Localization. *J Neurophysiol*
 567 22:613-636.
- 568 Goldstein MH, Kiang NYS (1958) Synchrony of neural activity in electric responses evoked by
 569 transient acoustic stimuli. *The Journal of the Acoustical Society of America* 30:107-114.
- 570 He W, Porsov E, Kemp D, Nuttall AL, Ren T (2012) The group delay and suppression pattern of the
 571 cochlear microphonic potential recorded at the round window. *PLoS One* 7:e34356.
- 572 Heinz MG, Colburn HS, Carney LHC (2001) Evaluating auditory performance limits: I. One-
 573 parameter discrimination using a computational model for the auditory nerve. *Neural*
 574 *Computation* 13:2273-2316.
- 575 Henry KR (1995) Auditory nerve neurophonic recorded from the round window of the Mongolian
 576 gerbil. *Hearing research* 90:176-184.
- 577 Javel E, Mott JB (1988) Physiological and psychophysical correlates of temporal processes in hearing.
 578 *Hearing research* 34:275-294.
- 579 Johnson DH (1980) The relationship between spike rate and synchrony in responses of auditory-nerve
 580 fibers to single tones. *J Acoust Soc Am* 68:1115-1122.
- 581 Joris PX, Verschooten E (2013) On the limit of neural phase locking to fine structure in humans. *Adv*
 582 *Exp Med Biol* 787:101-108.
- 583 Joris PX, Smith PH, Yin TC (1994a) Enhancement of neural synchronization in the anteroventral
 584 cochlear nucleus. II. Responses in the tuning curve tail. *Journal of neurophysiology* 71:1037-
 585 1051.
- 586 Joris PX, Carney LH, Smith PH, Yin TC (1994b) Enhancement of neural synchronization in the
 587 anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. *Journal of*
 588 *neurophysiology* 71:1022-1036.
- 589 Kiang NYS, Moxon EC, Kahn AR (1976) The Relationship of Gross Potentials Recorded from the
 590 Cochlea to Single Unit Activity in the Auditory Nerve. In: *Electrocochleography* (Ruben RJ,
 591 Elberling C, Salomon G, eds), pp 95-115. Baltimore: University Park Press.

- 592 Kiang NYS, Watanabe T, Thomas EC, Clark LF (1965) Discharge patterns of single fibers in the cat's
593 auditory nerve, 35 Edition. MIT Press, Cambridge: Research Monograph No 35.
- 594 Kidd RC, Weiss TF (1990) Mechanisms that degrade timing information in the cochlea. *Hear Res*
595 49:181-208.
- 596 Kim DO, Molnar CE (1979) A Population Study of Cochlear Nerve Fibers: Comparison of Spatial
597 Distributions of Average-Rate and Phase-Locking Measures of Responses to Single Tones. *J*
598 *Neurophysiol* 42:16-30.
- 599 Klumpp R, Eady H (1956) Some measurements of interaural time differences thresholds. *Journal of*
600 *the Acoustical Society of America* 28:859-864.
- 601 Lichtenhan JT, Cooper NP, Guinan JJ, Jr. (2013) A new auditory threshold estimation technique for
602 low frequencies: proof of concept. *Ear Hear* 34:42-51.
- 603 Lorenzi C, Gilbert G, Carn H, Garnier S, Moore BC (2006) Speech perception problems of the hearing
604 impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of*
605 *Sciences of the United States of America* 103:18866-18869.
- 606 McMahon CM, Patuzzi RB (2002) The origin of the 900 Hz spectral peak in spontaneous and sound-
607 evoked round-window electrical activity. *Hearing research* 173:134-152.
- 608 Moore BC (2008) The role of temporal fine structure processing in pitch perception, masking, and
609 speech perception for normal-hearing and hearing-impaired people. *J Assoc Res Otolaryngol*
610 9:399-406.
- 611 Palmer AR, Russell IJ (1986) Phase-locking in the cochlear nerve of the guinea-pig and its relation to
612 the receptor potential of inner hair cells. *Hear Res* 24:1-15.
- 613 Palmer AR, Shackleton TM (2009) Variation in the phase of response to low-frequency pure tones in
614 the guinea pig auditory nerve as functions of stimulus level and frequency. *Journal of the*
615 *Association for Research in Otolaryngology : JARO* 10:233-250.
- 616 Patuzzi RB, Brown DJ, McMahon CM, Halliday AF (2004) Determinants of the spectrum of the
617 neural electrical activity at the round window: transmitter release and neural depolarisation.
618 *Hearing Research* 190:87-108.

- 619 Prijs VF (1986) Single-unit response at the round window of the guinea pig. *Hearing research* 21:127-
620 133.
- 621 Rhode WS, Smith PH (1986) Encoding Timing and Intensity in the Ventral Cochlear Nucleus of the
622 Cat. *J Neurophysiol* 56:261-286.
- 623 Rose JE, Brugge JF, Anderson DJ, Hind JE (1967) Phase-locked response to low-frequency tones in
624 single auditory nerve fibers of the squirrel monkey. *J Neurophysiol* 30:769-793.
- 625 Rose JE, Brugge JF, Anderson DJ, Hind JE (1968) Patterns of activity in single auditory nerve fibres
626 of the squirrel monkey. In: de Reuck, AVS, Knight, J (Eds), *Ciba Foundation Symposium on*
627 *Hearing Mechanisms in Vertebrates* J&A Churchill, London, pp144–157.
- 628 Russell IJ, Sellick PM (1983) Low-frequency characteristics of intracellularly recorded receptor
629 potentials in guinea-pig cochlear hair cells. *J Physiol* 338:179-206.
- 630 Schiano JL, Trahiotis C, Bernstein LR (1986) Lateralization of low-frequency tones and narrow bands
631 of noise. *J Acoust Soc Am* 79:1563-1570.
- 632 Searchfield GD, Munoz DJ, Thorne PR (2004) Ensemble spontaneous activity in the guinea-pig
633 cochlear nerve. *Hear Res* 192:23-35.
- 634 Snyder RL, Schreiner CE (1984) The auditory neurophonic: basic properties. *Hearing research* 15:261-
635 280.
- 636 Starr A, Picton TW, Sininger Y, Hood LJ, Berlin CI (1996) Auditory neuropathy. *Brain* 119 (Pt
637 3):741-753.
- 638 Tasaki I (1954) Nerve impulses in individual auditory nerve fibers of guinea pig. *Journal of*
639 *Neurophysiology* 17:97-122.
- 640 Temchin AN, Ruggero MA (2010) Phase-locked responses to tones of chinchilla auditory nerve fibers:
641 implications for apical cochlear mechanics. *J Assoc Res Otolaryngol* 11:297-318.
- 642 van der Heijden M, Joris PX (2006) Panoramic measurements of the apex of the cochlea. *J Neurosci*
643 26:11462-11473.
- 644 Verschooten E, Joris PX (2014) Estimation of Neural Phase Locking from Stimulus-Evoked
645 Potentials. *J Assoc Res Otolaryngol*.

- 646 Versnel H, Prijs VF, Schoonhoven R (1992) Round-window recorded potential of single-fibre
647 discharge (unit response) in normal and noise-damaged cochleas. *Hearing research* 59:157-170.
- 648 Weiss TF, Rose C (1988a) Stages of degradation of timing information in the cochlea: A comparison
649 of hair-cell and nerve-fiber responses in the alligator lizard. *Hear Res* 33:167-174.
- 650 Weiss TF, Rose C (1988b) A comparison of synchronization filters in different auditory receptor
651 organs. *Hear Res* 33:175-180.
- 652 Zwislocki JJ, Feldman RS (1956) Just noticeable differences in dichotic phase. *J Acoust Soc Am*
653 28:860-864.

Figures

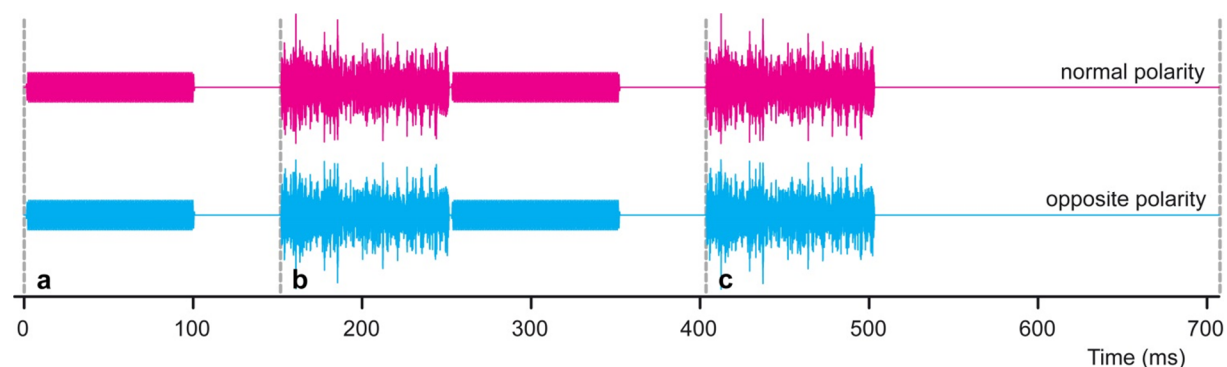


Figure 1 An illustration of the stimulus paradigm. It consists of two identical successive stimulus presentations (upper and lower), with opposite polarity. Each presentation has 3 segments, shown separated by vertical dashed lines: *a* contains the probe only, *b* the probe preceded by a masker, and *c* the masker only. The probe is a tone and the masker is a broadband noise (illustrated here) or a tone at the same frequency as the probe.

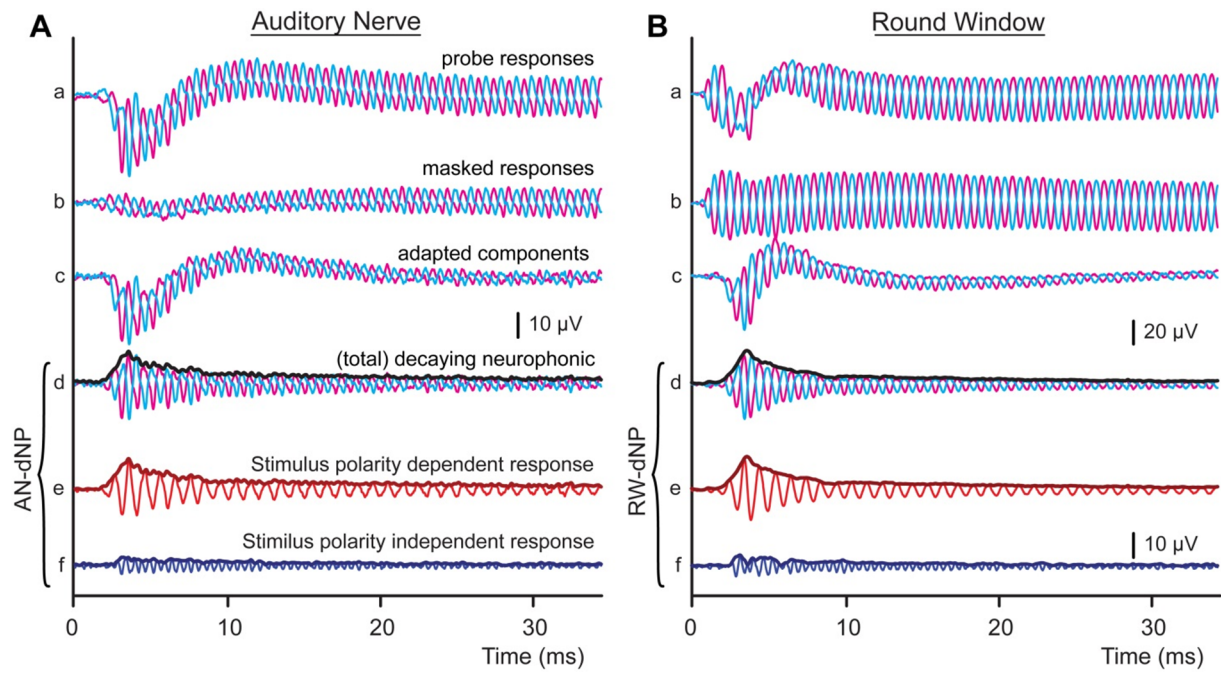
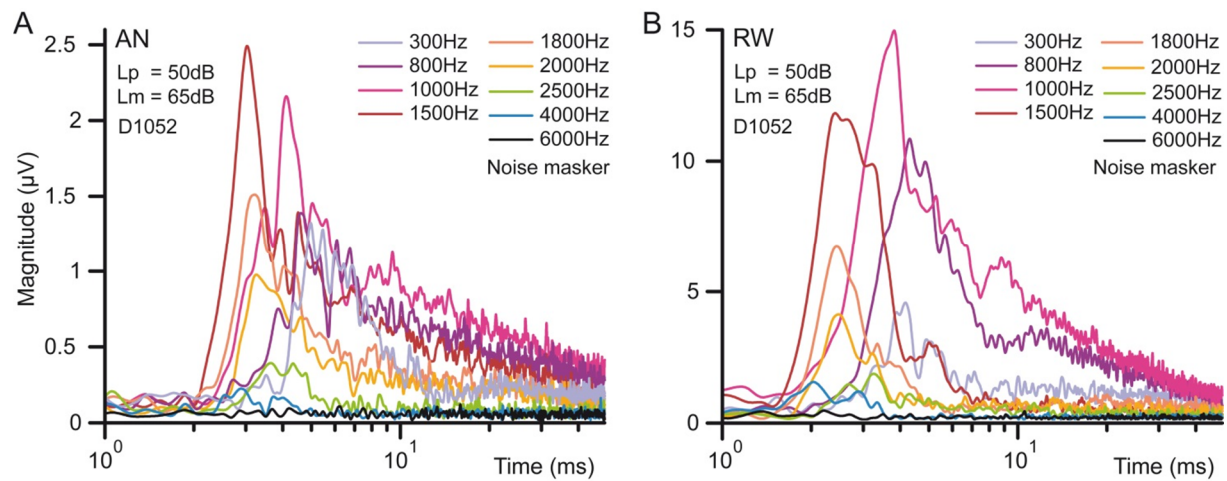


Figure 2 Evoked response averages from one animal for simultaneous recordings at the AN (**A**) and RW (**B**) for a probe tone of 1 kHz. Only the first 35 ms are shown. **Aa**, Pair of probe responses for two stimulus polarities. **Ab**, Same as **Aa**, but with a preceding tonal masker. **Ac**, Difference between signal **Aa** and **Ab**: the adapted component. **Ad**, Same as **Ac**, but with the CAP removed by high-pass filtering. **Ae** Halved difference of the responses shown in **Ad**. **Af**, Halved sum of the responses shown in **Ad**. **B**, similar to **A** but for RW. Probe parameters were: level = 50 dB SPL, frequency = 1 kHz, number of responses averaged = 128. Masker parameters were: level = 70 dB SPL, frequency = 1 kHz. For the final traces in each panel (**Ad-f**, **Bd-f**) the envelopes are shown as the thick line.

**Figure 3**

Effect of probe frequency on the time course of the magnitude of the recorded AN-dNP (**A**) and RW-dNP (**B**), simultaneously recorded in one animal. The probe tone level was 50 dB SPL and the masker was a broadband noise fixed at 65 dB SPL. To emphasize the early peak, time is plotted on a log-axis.

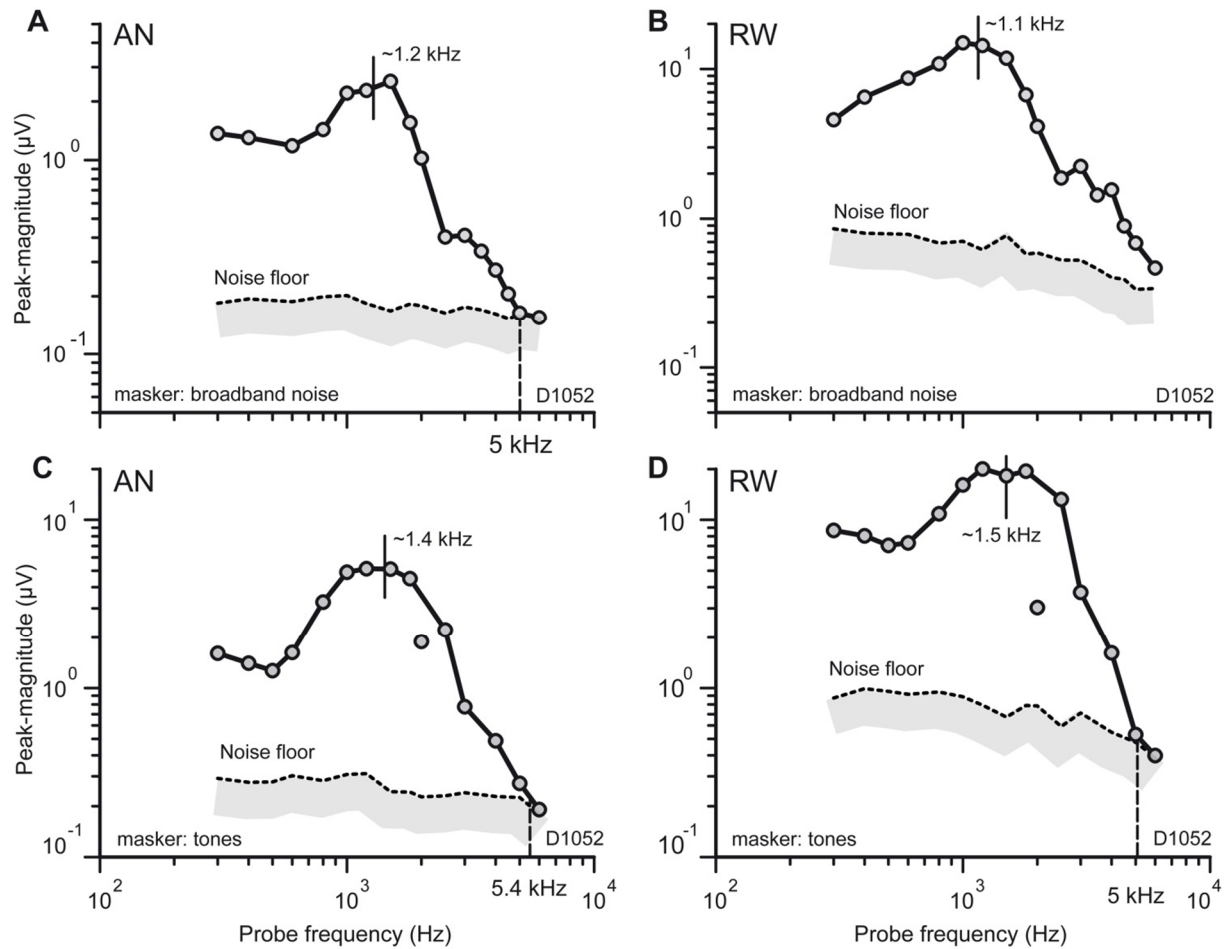


Figure 4 Peak-magnitudes of AN-dNP (*A,C*) and RW-dNP (*B,D*) as a function of probe frequency. The masker was either a broadband masker (50 Hz – 20 kHz, 65 dB SPL) (*A,B*) or a tonal masker of 65 dB SPL (*C,D*). The results are from the same animal as in Figure 3. The baseline is indicated by the dashed-shaded line and is the noise floor in the measurement. The estimated centers of the broad peaks in the curves are indicated by short vertical lines. The upper frequency limits are indicated by the dashed vertical lines, and are obtained at the intersection with the noise floor. The probe level was 55 dB SPL. Disconnected dots are considered to be outliers.

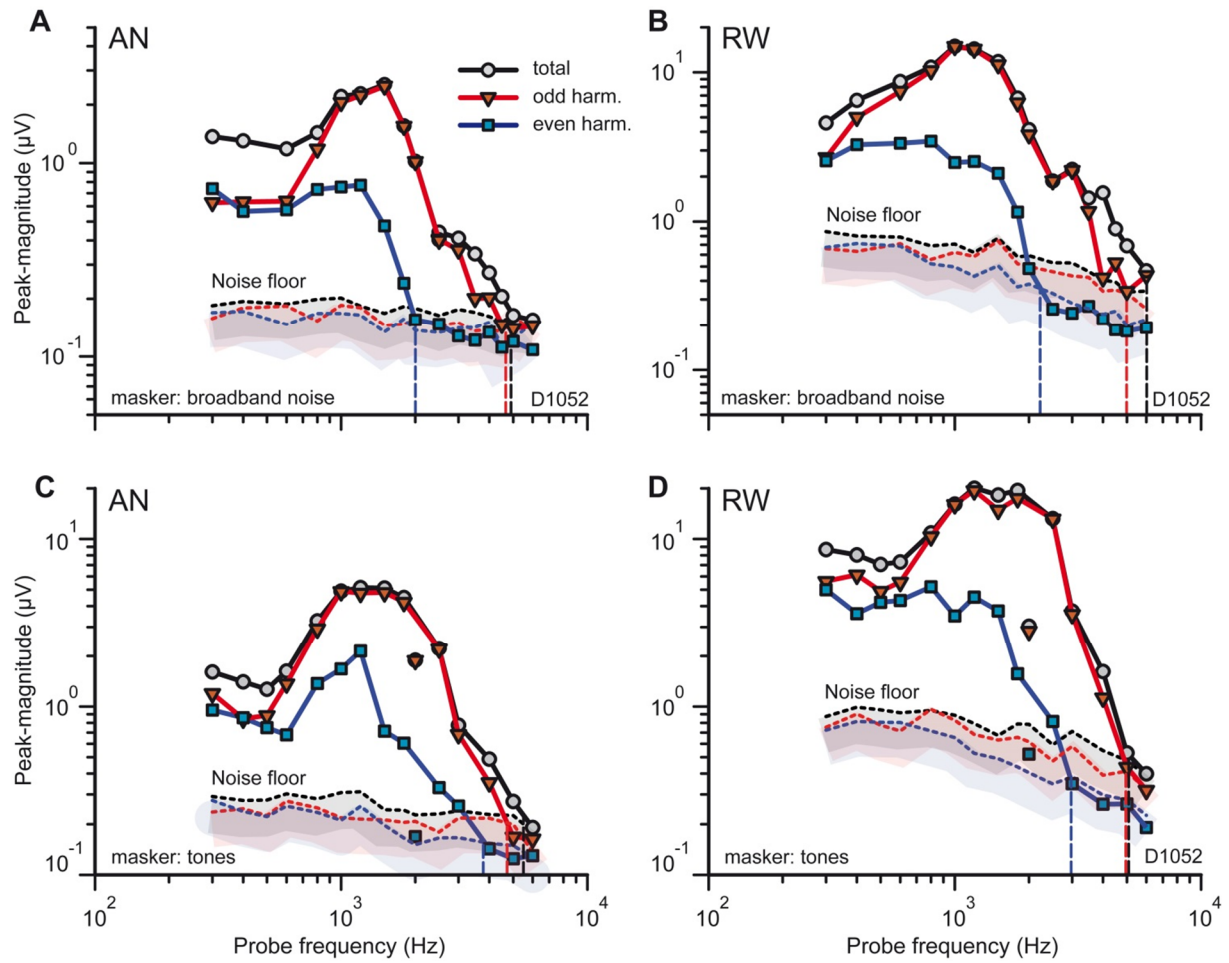


Figure 5 Peak-magnitudes of total maskable response (black, taken from Fig. 4) for broadband noise (A,B) and tonal maskers (C,D), with their stimulus polarity-dependent adapted components (odd harmonics; red filled triangles), and polarity-independent adapted components (even harmonics; blue filled squares). The noise floors of the measurements are indicated by the corresponding dashed and shaded lines. Disconnected symbols are considered outliers. Data from one animal.

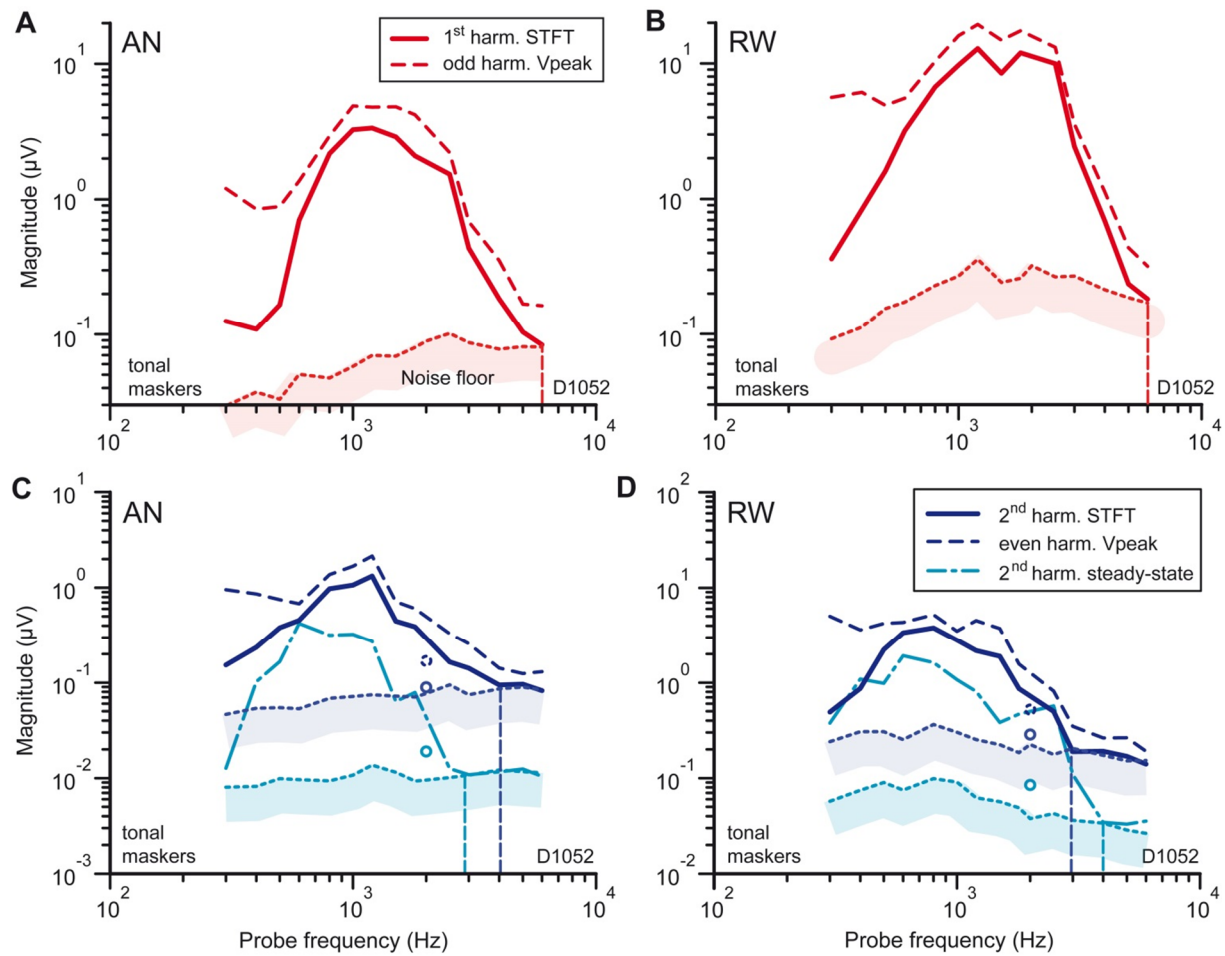


Figure 6 Comparison between peak-amplitudes of stimulus polarity dependent (**A,B**) and independent (**C,D**) and their 1st and 2nd harmonics obtained using the STFT-method applied to the dNP components obtained with tonal maskers, for the AN (**A,C**) and RW (**B,D**). For comparison with another measure for the neurophonic, sometimes called the ANOW, the 2nd harmonic steady-state amplitude (dashed dotted) is shown in **C** and **D**. The dots are considered to be outliers. Data from one animal

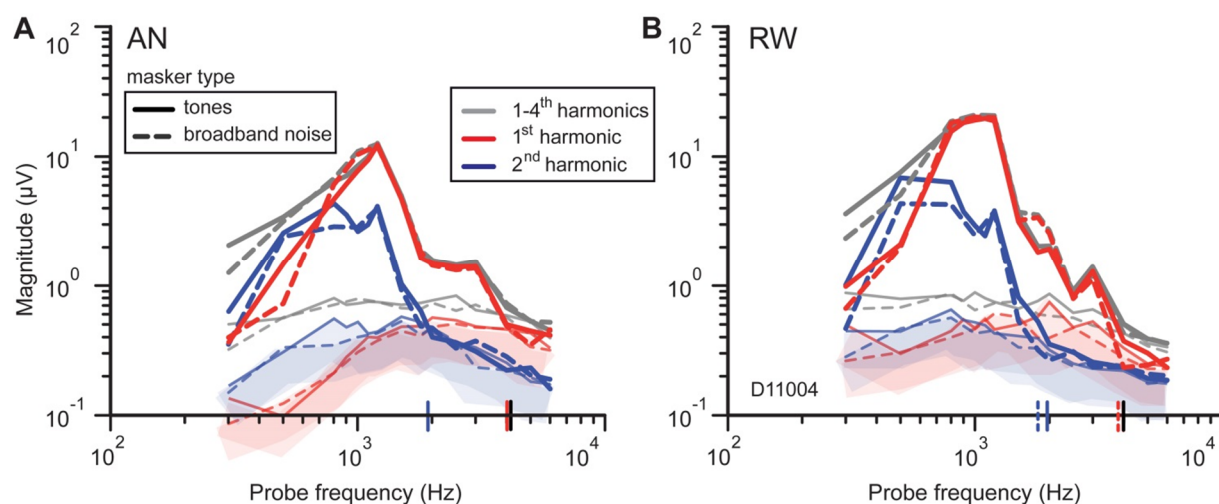


Figure 7 Comparison between the AN-dNP (*A*) and RW-dNP (*B*) magnitudes for tonal (solid lines) and broadband noise (dashed lines) maskers obtained with the STFT from one animal. The panels contain maximum amplitudes obtained from the first four harmonic contributions (gray); the 1st harmonic (red); and the 2nd harmonic (blue). The level for the probes was 55 dB SPL and for both type of maskers the level was 70 dB SPL. The corresponding noise floors are indicated by thin lines and shading. The vertical lines crossing the abscissa indicate the upper limits.

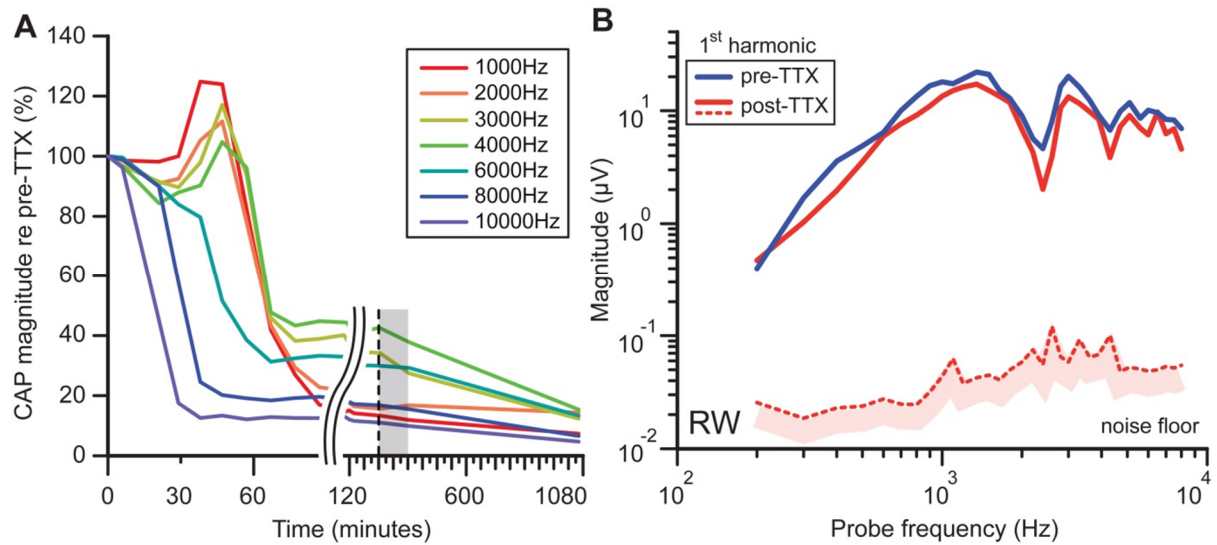


Figure 8 Reduction of CAP but not CM after TTX application in one animal. **A**, CAP reduction as a function of time after administration of TTX in the niche of the RW. The parameter is the probe frequency. The vertical dashed line indicates the start of the post-TTX recordings; the gray zone covers the time period of the measurement. **B**, Comparison between 1st harmonic RW-CM magnitudes before (blue line, taken at time 0 in panel **A**) and after (solid red line, taken during gray zone in panel **A**) TTX administration. The background noise of the post-TTX RW-CM is included as the red dotted line and shading

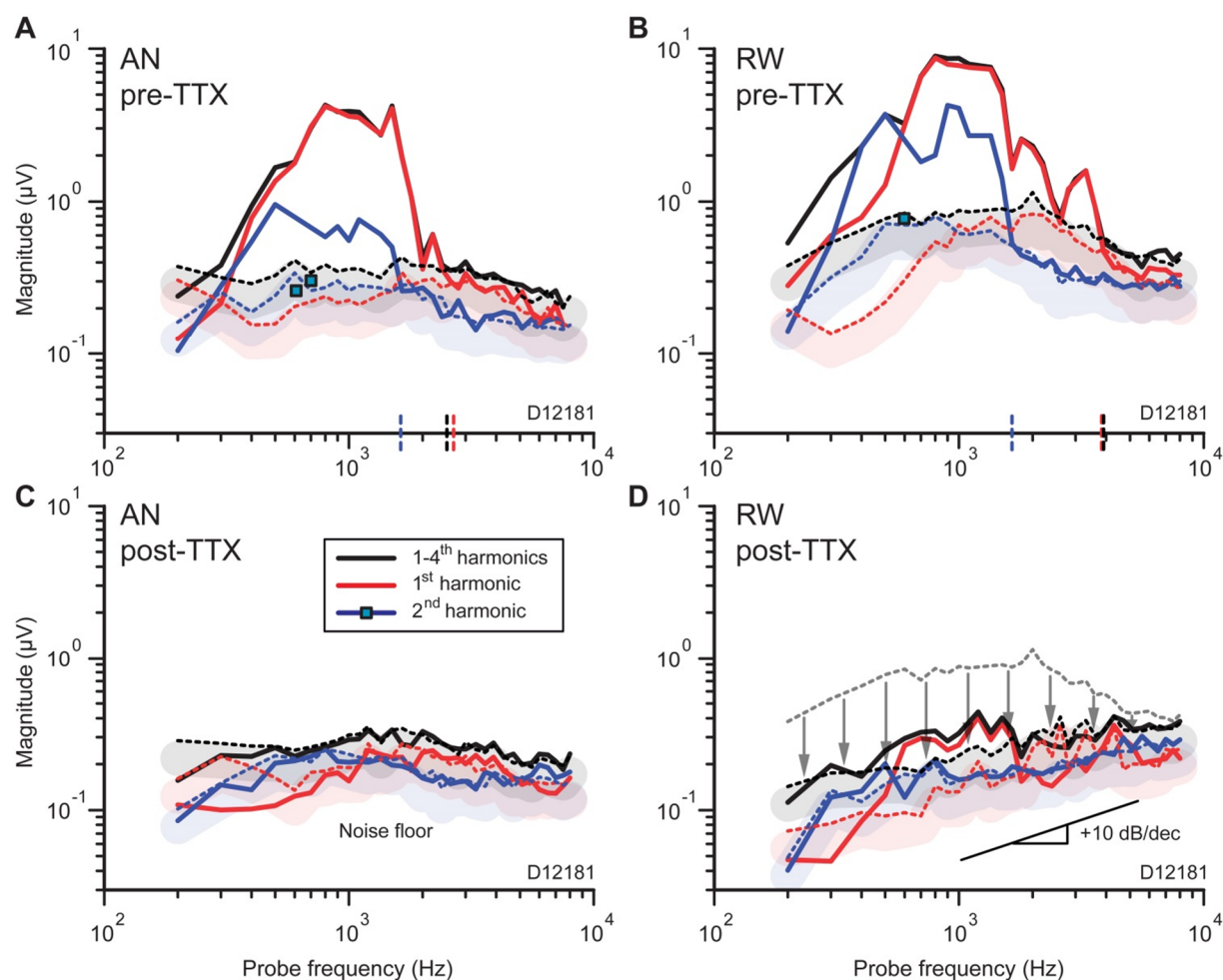


Figure 9 The magnitude of the AN-dNP and RW-dNP before (**A,B**) and after (**C,D**) the administration of TTX, for the AN (**A,C**) and RW (**B,D**). The corresponding noise floors are indicated by the dotted lines. The slope of +10 dB/decade in **D** indicates the slope imposed by the frequency proportional filter in the STFT method. The gray arrows in panel **D** indicate the reduction of the (averaged) background noise at the RW as a result of the administration of TTX. Data from one animal.

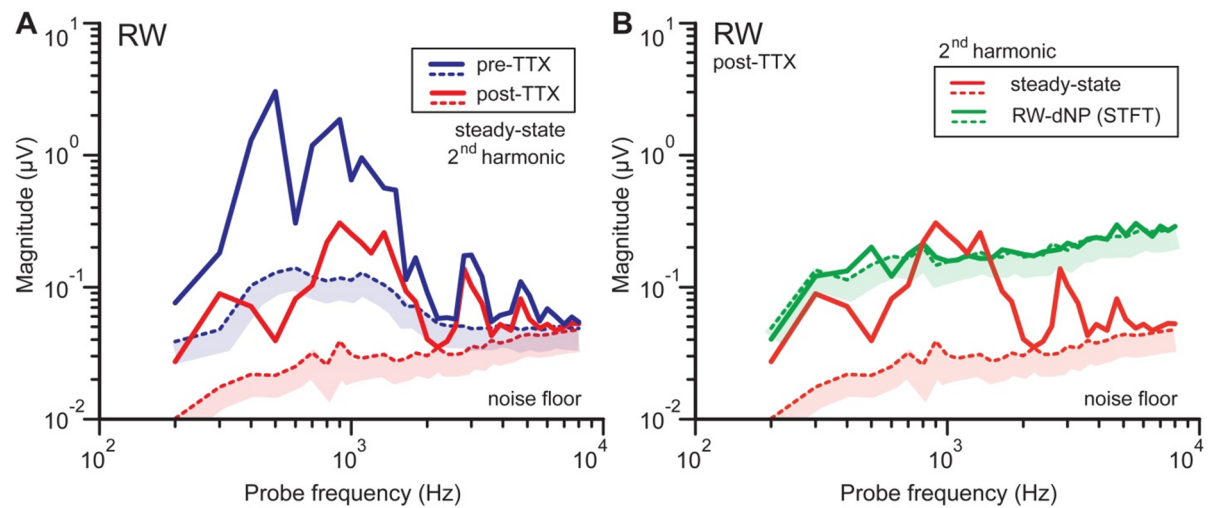


Figure 10 Comparison of the post-TTX 2nd harmonic steady-state amplitude (red) with **A**, the pre-TTX amplitude (blue) and **B**, the 2nd harmonic STFT magnitudes (green). The amplitudes were obtained from the steady-state part of the probe response using an FFT with long time window.

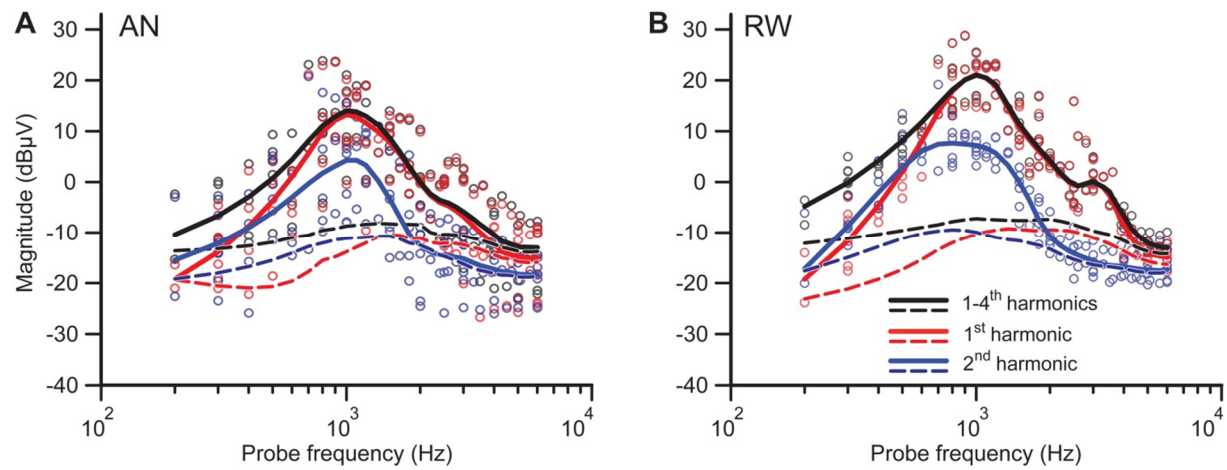


Figure 11 General mean trends for the different harmonics of the neurophonic using the STFT method for the AN (**A**) and RW (**B**). The trends (MATLAB, *RLOESS*, span: 0.35) were obtained from 6 measurement sessions in 4 animals. The corresponding noise floors are indicated by dashed lines. Note that magnitude is in dB re 1 μ V_{RMS}.

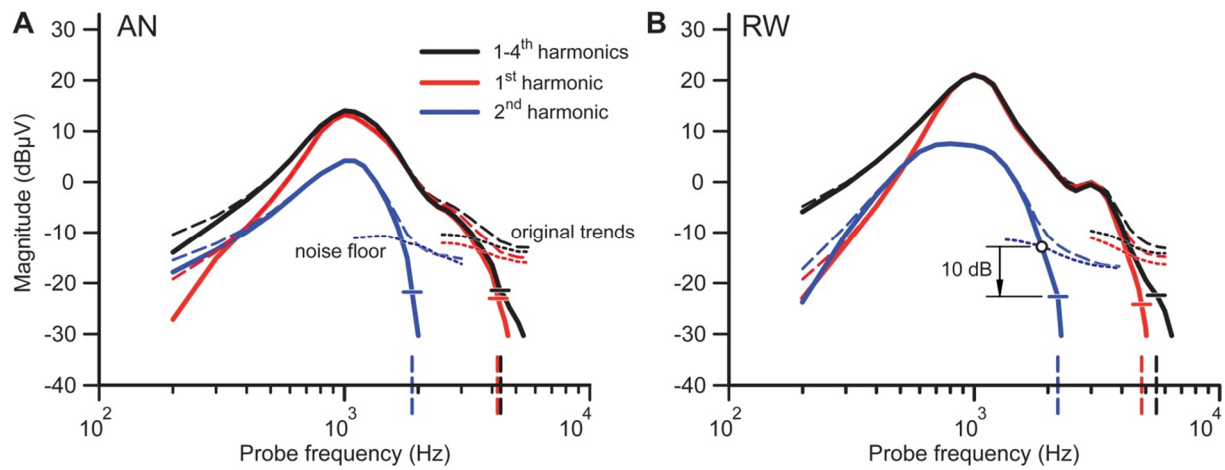


Figure 12 Similar to Figure 11 (dashed lines), but compensated (solid lines) for the noise floor. The upper frequency limits are obtained at the intersection of the compensated trend lines with the horizontal small lines, which are the magnitude obtained at 10 dB below the intersection of the noise floor with the compensated curve. The corresponding noise floors are indicated by dotted lines. The upper frequency limit is indicated with vertical dashed lines on the abscissa.

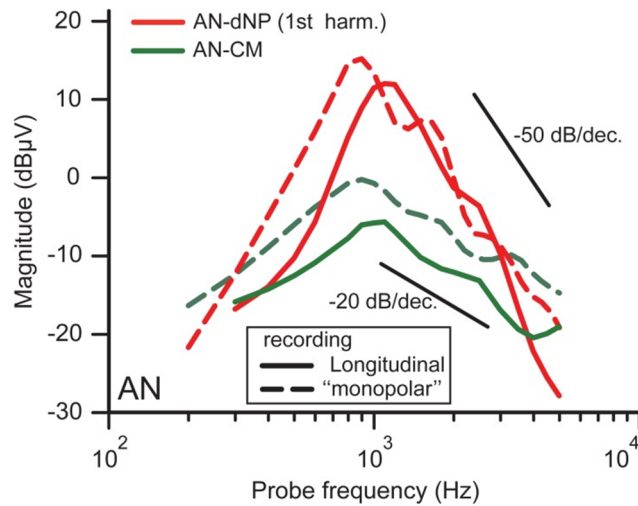


Figure 13 Mean trends of the AN-CM and AN-dNP for two different recording configurations.

In the longitudinal configuration the electrodes are placed along the AN; the positive electrode was positioned near the internal auditory meatus and the reference electrode at the merging of the AN with the CN. In the "monopolar" method the positive electrode was positioned on the nerve and the reference electrode in the nape of the neck. The AN-dNP data were obtained with the STFT method; the AN-CM data were obtained from the masked signal at neural onset. The black lines indicate roll-offs for reference. The results for the different recording configurations are from 2 different cats.

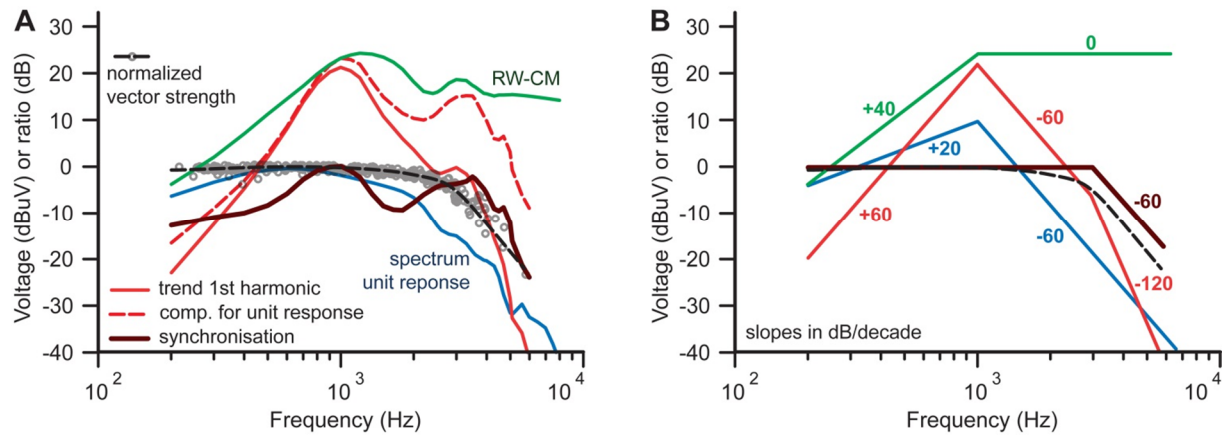


Figure 14 Synchronization index for the neurophonic. (**A**) Different frequency dependent curves are shown: the spectrum of mean unit contributions (blue, from 4 examples in Kiang et al. (1976)); the mean trend of the RW-CM (green); the mean trends of the 1st harmonic component of the neurophonic (Fig. 12B, red curve) and its compensation for the filtering effect of the waveform of the unit response (red dashed curve); the synchronization index for the neurophonic (SIN, brown line; calculated as the ratio between the red dashed curve and green curve); and the VS (expressed in dB) for single AN-fibers in cat (black dashed line; the single data points indicates as gray circles are from Figure 5 of Johnson (1980)). **B**, Same as in **A** but with idealized curves. The slopes (dB/decade) are indicated next to the curves.